

Washington University School of Medicine Digital Commons@Becker

Open Access Publications

2007

UV-B responsive microRNA genes in *Arabidopsis thaliana*

Xuefeng Zhou

Washington University in St Louis

Guandong Wang

Washington University in St Louis

Weixiong Zhang

Washington University School of Medicine in St. Louis

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation

Zhou, Xuefeng; Wang, Guandong; and Zhang, Weixiong, "UV-B responsive microRNA genes in *Arabidopsis thaliana*." *Molecular Systems Biology*.3,. Article number: 103. (2007).
http://digitalcommons.wustl.edu/open_access_pubs/1369

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.

REPORT

UV-B responsive microRNA genes in *Arabidopsis thaliana*

Xuefeng Zhou¹, Guandong Wang¹ and Weixiong Zhang^{1,2,*}

¹ Department of Computer Science and Engineering, Washington University in Saint Louis, Saint Louis, MO, USA and ² Department of Genetics, Washington University in Saint Louis, Saint Louis, MO, USA

* Corresponding author. Department of Computer Science and Engineering, Department of Genetics, Washington University in Saint Louis, Saint Louis, MO 63130-4899, USA. Tel.: +1 314 935 8788; Fax: +1 314 935 7302; E-mail: zhang@cse.wustl.edu

All authors contributed equally to this work

Received 16.1.07; accepted 11.2.07

MicroRNAs (miRNAs) are small, non-coding RNAs that play critical roles in post-transcriptional gene regulation. In plants, mature miRNAs pair with complementary sites on mRNAs and subsequently lead to cleavage and degradation of the mRNAs. Many miRNAs target mRNAs that encode transcription factors; therefore, they regulate the expression of many downstream genes. In this study, we carry out a survey of *Arabidopsis* microRNA genes in response to UV-B radiation, an important adverse abiotic stress. We develop a novel computational approach to identify microRNA genes induced by UV-B radiation and characterize their functions in regulating gene expression. We report that in *A. thaliana*, 21 microRNA genes in 11 microRNA families are upregulated under UV-B stress condition. We also discuss putative transcriptional downregulation pathways triggered by the induction of these microRNA genes. Moreover, our approach can be directly applied to miRNAs responding to other abiotic and biotic stresses and extended to miRNAs in other plants and metazoans.

Molecular Systems Biology 17 April 2007; doi:10.1038/msb4100143

Subject Categories: computational methods; RNA

Keywords: integration of heterogeneous data; microRNA gene; UV-B responsive

Introduction

MicroRNAs (miRNAs) are approximately 22-nucleotide-long, non-coding RNAs that play critical roles in regulating gene expression at the post-transcriptional level (Bartel, 2004; He and Hannon, 2004). The discovery of miRNAs has broadened our perspectives on the mechanisms of repression of gene expression, which is an important regulatory mechanism mediating many biological processes such as development, cell proliferation and differentiation. In plants, mature miRNAs base-pair with complementary sites on target mRNAs and subsequently direct the mRNAs to be cleaved or degraded. Plant miRNAs regulate many genes that are involved in developmental control, for example, auxin signaling (Bonnet *et al*, 2004; Jones-Rhoades and Bartel, 2004), organ polarity (Eshed *et al*, 2001; McConnell *et al*, 2001; Kidner and Martienssen, 2004), development transitions (Aukerman and Sakai, 2003; Chen, 2004), leaf growth (Palatnik *et al*, 2003) and RNA metabolism (Xie *et al*, 2003; Vaucheret *et al*, 2004). Several recent studies showed important functions of miRNAs in response to adverse abiotic stresses (Jones-Rhoades and Bartel, 2004; Sunkar and Zhu, 2004; Lu *et al*, 2005; Bari *et al*, 2006). In *Arabidopsis*, *miR399* was identified to be highly

expressed under phosphate starvation (Fujii *et al*, 2005; Bari *et al*, 2006; Chiou *et al*, 2006) and *miR395* was identified to be induced under sulfate starvation (Jones-Rhoades and Bartel, 2004). Furthermore, quantitative experimental analysis proved that *miR393* was strongly induced under cold stress (Sunkar and Zhu, 2004). In *Populus*, moreover, some miRNAs can be induced by mechanical stress and may function in critical defense systems for structural and mechanical fitness (Lu *et al*, 2005).

Many targets of miRNAs encode transcription factors as well (Bartel, 2004), each of which further regulates a set of downstream genes. Thus, activation of miRNA genes under abiotic stresses will lead to the repression of many downstream protein coding genes and affect physiological responses. Among various environmental factors, light plays a particularly important role. Sunlight is not only the energy source for plant photosynthesis, but also regulates several plant developmental processes and some physiological processes such as photosynthesis, seasonal and diurnal time sensing (Chory *et al*, 1996; Chattopadhyay *et al*, 1998; Baroli *et al*, 2004; Jiao *et al*, 2005). Similar to adverse environmental factors such as drought and salinity, light can also have stress effect on plants. It interacts with endogenous developmental

programs, hence affects plant growth and development. In order to acclimate under such conditions, specific photoreceptor systems have been developed and evolved to monitor changes of light composition (Dunaeva and Adamska, 2001; Harvaux and Kloppstech, 2001; Shao *et al*, 2006). With complex photoreceptors, plant can register UV-B radiation and transduce the information to nucleus, hence affect gene expression (Chattopadhyay *et al*, 1998; Kimura *et al*, 2003b; Jiao *et al*, 2005). Changes in gene expression in response to UV-B radiation include reduction in expression and synthesis of key photosynthetic proteins as well as perturbation of expressions of the genes involved in defense mechanisms (Chattopadhyay *et al*, 1998; Kimura *et al*, 2003b; Jiao *et al*, 2005).

Regulation of gene expression plays an important role in a variety of biological processes such as development and responses to environmental stimuli. In plants, transcriptional regulation is mediated by a large number of transcription factors (TFs) controlling the expression of tens or hundreds of target genes in various, sometimes intertwined, signal transduction cascades (Venter and Botha, 2004; Wellmer and Riechmann, 2005). Transcription factor binding sites (TFBSs) are functional short DNA sequences (*cis*-elements) that determine the timing and location of transcriptional activities. Many computational methods have been developed to reveal relationships between gene expression patterns and TFBSs in the proximal upstream regulatory regions of the genes of interest. In yeast, motifs with known functions have been related to transcriptional pathways by statistical analysis of the occurrence of known motifs in the promoters of coregulated genes (Bussemaker *et al*, 2001). However, the presence of individual motifs is only marginally indicative of a gene's expression pattern. Extended strategies pursue to optimally predict gene expression patterns with promoter *cis*-elements and their combinations. With a systematic strategy, the expression of a large proportion of genes in *Saccharomyces cerevisiae* was accurately predicted based on promoter sequences (Beer and Tavazoie, 2004). Although distal regulatory elements other than those in proximal upstream promoter regions can modulate gene expression, a recent study emphasized that the sequences in the 5'-upstream regions of genes were of primary importance in *Arabidopsis* gene regulation (Lee *et al*, 2006). Specifically, promoter sequences were sufficient to recapture mRNA expression levels for 80% of the TFs studied. This study confirmed the important role of promoter regions in *Arabidopsis* gene expression.

In light of this transcriptome-based perspective and by taking advantage of the vast available data of genome-scale microarray expression profile of protein-coding genes, we develop an innovative computational approach to explore the expression activity of miRNA genes under certain conditions. We focus on identifying and annotating miRNAs in *A. thaliana*, which are responsive to UV-B radiation, and further consider the regulatory pathways that are probably affected by the putative UV-B-inducible miRNA genes. Our approach is based on the following two observations. First, plant miRNAs generally direct endonucleolytic cleavage of target mRNAs (Llave *et al*, 2002; Schwab *et al*, 2005), hence enable rapid clearance of target mRNAs when they are expressed (Bartel, 2004; Axtell and Bartel, 2005). Under a particular condition, if

an miRNA is upregulated, its targets are most likely to be coherently downregulated. Second, miRNA genes are transcribed by RNA polymerase II (Lee *et al*, 2004; Houbaviy *et al*, 2005; Xie *et al*, 2005; Zhou *et al*, 2007). Hence, the 5' proximal promoters of miRNA genes are the most important regulatory regions, and significant *cis*-elements in these regions are important in determining the spatial and temporal expression patterns of the miRNA genes. Therefore, miRNA and protein-coding genes carrying the same or similar *cis*-elements in their promoters are very likely to be coregulated under the same condition and consequently very likely to be coexpressed.

Although we focus on *Arabidopsis* UV-B responding miRNA genes in this study, our approach can be directly applied to plant miRNA genes functioning under other abiotic or biotic stress conditions.

Results and discussions

UV-B responsive miRNAs

One of the bases of our method for finding stress responsive miRNA genes is that protein-coding genes targeted by the same miRNA are likely to have coherently downregulated expression patterns. We consider an miRNA to be putatively stress inducible if the expressions of its target genes are coherently repressed and the coherence is statistically significant above a threshold. In this study, we only considered bonafide target genes reported in the literature. For each miRNA, pairwise cosine similarities of the expressions of its target genes were computed. We measured the coherence of the expressions of its target genes by average pairwise similarity. Statistical significance of the coherence was assessed with a *P*-value from a Monte Carlo simulation. Briefly, for each miRNA with *n* target genes, we first calculated the average pairwise cosine similarity of the expressions of the target genes. We then randomly sampled *n* genes from the whole set of genes that were profiled and calculated their average pairwise cosine similarity. We repeated the sampling a large number of times, for instance, a million times in our study, and took as an empirical *P*-value the frequency of observing a similarity value larger than that of the target genes. For each miRNA, we repeated this simulation 100 times and calculated the average *P*-value and standard deviation.

Table 1 Putative UV-B responsive miRNAs

Gene id	No. of targets	Cosine similarity	<i>P</i> -value	s.d. ^a
miR156/157	10	0.28	1.09E-2	2.43E-4
miR159/319	12	0.27	7.11E-3	4.30E-5
miR160	3	0.63	1.42E-2	1.94E-4
miR165/166	5	0.46	1.17E-2	1.11E-4
miR167	2	0.64	6.10E-2	3.47E-4
miR169	7	0.66	3.60E-5	7.00E-6
miR170/171	3	0.61	1.69E-2	1.61E-4
miR172	5	0.66	4.25E-4	1.70E-5
miR393	6	0.28	4.33E-2	2.67E-4
miR398	3	0.84	2.44E-4	6.00E-5
miR401	2	0.73	2.96E-2	2.71E-4

^aStandard deviations of *P*-values.

Table I shows putative UV-B responsive miRNAs. For each of these miRNAs, its target genes are coherently downregulated, and the coherence of their expression patterns is statistically significant. Except *miR167*, whose *P*-value is less than 0.07, all candidates have *P*-values less than 0.05.

For *miR158*, *miR162*, *miR163*, *miR168*, *miR395*, *miR402*, *miR403*, *miR404*, *miR405* and *miR406*, we only found one bona fide target gene in the microarray data set, and could not test their coherence, and thus excluded them from our study.

UV-B responsive miRNA genes

We applied our computational approach, discussed in section outline of the computational approach, to the microarray gene expression data under UV-B radiation treatment from the *Arabidopsis* AtGenExpress project (www.arabidopsis.org/info/expression/ATGenExpress.jsp). We predicted 21 miRNA genes in 11 miRNA families to be upregulated under UV-B radiation. Table II lists these UV-B-inducible miRNA genes. A putative UV-B responsive miRNA gene must satisfy two criteria: First, the set of protein-coding genes with the same array of motifs in their proximal promoter regions is enriched with UV-B upregulated genes. Second, its inferred expression (discussed below) should be anticorrelated with the expressions of its target genes.

For each miRNA gene, we analyzed whether the combination of significant motifs in its promoter was statistically relevant to the UV-B stress. First, we examined all protein-coding genes in the whole set of gene profiled in the microarray experiments, and found those genes that contain the same or very similar motifs in their proximal promoter regions.

Table II Putative UV-B responsive miRNA genes

Gene id	<i>P</i> -value ^a	Cosine similarity	<i>P</i> -value ^b	s.d. ^c
miR156b	2.15E-8	-0.42	5.11E-2	2.21E-4
miR156e	2.48E-5	-0.42	5.81E-2	1.76E-4
miR156f	9.39E-2	-0.42	2.69E-2	1.05E-4
miR156h	1.67E-6	-0.41	5.76E-2	2.73E-4
miR157c	8.93E-2	-0.32	8.32E-2	2.76E-4
miR159a	8.51E-2	-0.41	3.09E-2	1.95E-4
miR159b	9.94E-4	-0.48	1.25E-2	1.11E-4
miR160c	3.48E-4	-0.53	7.14E-2	3.51E-4
miR165a	2.11E-11	-0.52	5.24E-2	2.10E-4
miR166c	4.55E-2	-0.47	7.50E-2	1.60E-4
miR166f	2.50E-7	-0.47	9.45E-2	3.64E-4
miR167d	2.52E-6	-0.72	7.81E-2	1.77E-4
miR169d	1.79E-2	-0.41	9.10E-2	3.21E-4
miR169j	5.36E-10	-0.41	9.37E-2	1.93E-4
miR170	1.27E-2	-0.69	2.82E-2	1.83E-4
miR171a	4.66E-2	-0.75	8.32E-3	9.10E-5
miR172c	1.16E-2	-0.75	1.15E-3	4.20E-5
miR172e	8.21E-5	-0.77	2.60E-4	1.80E-5
miR393a	1.08E-5	-0.60	4.52E-2	1.66E-4
miR398a	4.89E-2	-0.78	5.94E-2	2.13E-4
miR401	1.2E-12	-0.71	8.51E-2	1.97E-4
miR168a	7.2E-5			
miR395c	8.09E-2			
miR395e	8.63E-3			

^a*P*-values for assessing the enrichment of UVB upregulated genes in the set of coding genes that contain the same arrays of motifs as the miRNA genes.

^b*P*-values for assessing the significance of the cosine similarities.

^cStandard deviations of the *P*-values^a.

We then tested whether these protein-coding genes were enriched with upregulated genes (see sections outline of the computational approach and hypergeometric distribution of motifs or motif combinations).

We further imposed on miRNA genes a criterion of anticorrelation between the inferred expression of an miRNA gene and the expressions of its mRNA targets, in order to filter out possible false predictions. As we did not rely on any direct information of miRNA expression, we used the inferred expression of an miRNA gene and the expressions of its targets to compute their anticorrelation (see section inference of expression patterns of miRNA genes). In our study, we chose the five best protein-coding genes whose 5' proximal promoters contain arrays of *cis*-elements that most resemble those of the corresponding miRNA genes. These five genes are most likely to be coregulated with the corresponding miRNA gene, and thus their expression patterns are most likely to be similar to the expression pattern of the miRNA gene. In the rest of our discussion, we refer to the average expression pattern of the top five coregulated protein-coding genes of an miRNA as its inferred expression pattern or expression pattern for short.

Before we inferred expression patterns of miRNA genes, we applied the inferring procedure to 100 randomly selected protein-coding genes with known expression patterns, and then assessed similarities between their inferred and actual expression patterns. For all these 100 genes, the cosine similarity values of their inferred and actual expression patterns are between 0.3 and 0.89, and the average of these values is 0.51. Figure 1 shows the inferred expression pattern and the actual expression pattern of protein-coding gene, *At1g19770*. The figure gives a pictorial view of the similarity of the inferred and original expression patterns. The cosine similarity value of these two expression patterns of *At1g19770* is 0.76.

For each putative UV-B responsive miRNA gene, we calculated the average cosine similarity between its inferred expression and the expressions of its targets. We assessed the

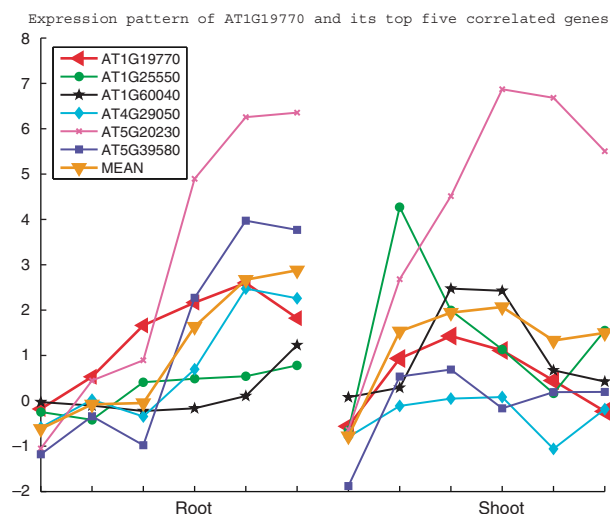


Figure 1 The expressions of *At1g19770* in root and shoot, respectively, the expressions of the five protein-coding genes that are most correlated to it and their mean expression.

statistical significance of the similarity with a *P*-value. Similar to the analysis of expression coherence of target genes in the section UV-B responsive miRNAs, *P*-value was also obtained by a Monte Carlo simulation. We took as an empirical *P*-value the frequency to observe a cosine similarity value smaller than that in the real data. For each miRNA, the simulation was also repeated 100 times to obtain an average *P*-value and a standard deviation.

Forty miRNA genes satisfy the first criterion. However, as shown in Table II, inferred expressions of 21 miRNA genes are anticorrelated to the expressions of their target genes (cosine similarity less than 0), and the anticorrelations reflected by the average cosine similarities of inferred expressions and expressions of target genes are statistically significant. These 21 genes are our predicted UV-B responsive miRNA genes.

In all putative UV-B responsive miRNA families shown in Table I, at least one member gene from each family was predicted to be upregulated under UV-B radiation. However, none of the members in other miRNA families was predicted to be UV-B responsive. Three miRNA genes, *miR168a*, *miR395c* and *miR395e*, might also be UV-B responsive. The arrays of motifs in their proximal promoter regions are statistically significantly relevant to UV-B stress, shown by small *P*-values obtained from an accumulative hypergeometric test. Protein-coding genes sharing the same array of motifs with them have enriched GO (gene ontology) terms that are related to stress response (see discussion in the section functions of protein-coding genes that share the same arrays of motifs as putative UV-B responsive miRNA genes). However, as these miRNAs have fewer than two experimentally validated target genes, the coherence of their target gene expressions and the anticorrelations between their expressions and expressions of their targets could not be analyzed. Hence these genes will not be discussed further.

Functions of protein-coding genes that share the same arrays of motifs as putative UV-B responsive miRNA genes

For each putative UV-B responsive miRNA gene shown in Table II, there are some protein-coding genes containing in their proximal promoter regions the same array of motifs. These protein-coding genes are very likely to share the same regulatory program, hence coexpress with the miRNA gene. In order to further interpret their relevance to UV-B stress, and hence to confirm the relevance of the miRNA gene to UV-B stress, we calculated the enrichment of GO functional terms in the annotations of these protein-coding genes. As shown in Table III, for 13 out of 21 putative UV-B responsive miRNA genes, we identified significantly enriched stress-related GO terms, by using the webserver for gene annotation analysis, FuncAssociate (llama.med.harvard.edu/cgi/func/funcassociate). In the table, the *P*-values represent statistical significance of the GO terms.

These enriched GO terms can be grouped into three major categories. The first category is related to transcription regulation. Protein-coding genes sharing the same regulatory regions as four miRNA genes, *miR156b*, *miR165a*, *miR169j* and *miR172c*, fall into this category. The second category is

related to direct response to stress or external stimuli. Protein-coding genes corresponding to *miR156h* and *miR166f* are in this category. The last category includes hydrolase activity and oxidoreductase activity. Protein-coding genes that are likely regulated by the same regulatory programs as the rest seven miRNA genes are in this category. It has been well studied that hydrolase and oxidoreductase are involved in response to many stresses, including light stress (Kimura *et al*, 2001, 2003a; Apel and Hirt, 2004). The analysis of GO term enrichment provides additional evidence that these 13 miRNA genes are very likely to be involved in the responses to the UV-B stress. Three miRNA genes, *miR168a*, *miR395c* and *miR395e*, were excluded from our prediction due to lack of reported target genes. However, protein-coding genes containing the same arrays of motifs as these three miRNA genes, especially *miR168a* and *miR395c*, are enriched with stress-related GO terms.

Presence of known light-relevant *cis*-elements in the promoters of miRNA genes

Using the WordSpy genome-wide motif finding algorithm (Wang *et al*, 2005; Wang and Zhang, 2006), we identified many significant *cis*-elements that are characteristic of UV-B responsiveness of the 21 miRNA genes listed in Table II. Some of them are well characterized in plant motif databases such as PLACE (Higo *et al*, 1999) and discussed in the literature (Terzaghi and Cashmore, 1995; Narusaka *et al*, 2004; Zhang *et al*, 2004; Yamaguchi-Shinozaki and Shinozaki, 2005). The *cis*-elements, which have been experimentally characterized in light-regulated genes, include the G-box (CACGTG), the GT-1 site (GGTAA), I-boxes (GATAAGA), TGA-box (TGACGT), GATA-box (GATATTT), H-box (CCTACC) and CCAAT (CCAAT) (Terzaghi and Cashmore, 1995). As shown in Table IV, these motifs appear in the promoters of some of miRNA genes that are upregulated by UV-B stimuli.

The presence of the well-studied light-related motifs shed light on the possible mechanisms activating the miRNA genes. For these 21 miRNAs, the most prevalent *cis*-elements are GT-1 site, I-box core and CCAAT-box. These miRNA genes all have GT-1 site in their promoters, all except one contain I-box core and 17 of them contain CCAAT-box. The involvement of GT-1 site, I-box and CCAAT-box in abiotic stress regulation has been well studied (Teakle and Kay, 1995; Arguello-Astorga and Herrera-Estrella, 1998; Shinozaki *et al*, 2003; Zhang *et al*, 2004; Yamaguchi-Shinozaki and Shinozaki, 2005); therefore, it is not surprising to find them in almost all of these UV-B-induced miRNA genes. Among the 21 miRNA genes, seven contain the GATA-box in their promoters, which has been shown to regulate light responsive genes (Teakle and Kay, 1995; Arguello-Astorga and Herrera-Estrella, 1998).

These motifs were previously analyzed on protein-coding genes. Their presence suggests that these miRNA genes are regulated similarly as light responsive protein-coding genes. To be specific, we list in Table V the known motifs that *miR167d* shares with its possible coregulated protein-coding genes. These shared motifs provide additional evidence that these miRNA genes and the corresponding protein-coding genes are regulated by similar mechanisms under UV-B stimuli.

Table III Stress-related GO terms are enriched in the annotations of protein-coding genes that contain all the motifs present in corresponding putative UV-B responsive miRNA genes

Gene id	P-value	GO term
miR156b	1.6E-5	Membrane
	2.2E-5	Binding/ligand
	2.3E-5	Transcription factor activity
	2.4E-5	Transcription regular activity
	3.1E-5	Hydrolase activity, acting on glycosyl bonds/N-glycosylase/glycosylase
MiR156e	3.9E-5	Catalytic activity/enzyme activity
	3.2E-6	Hydrolase activity, acting on glycosyl bonds/N-glycosylase/glycosylase
	1.4E-5	Hydrolase activity, hydrolyzing O-glycosyl compounds/O-glucosyl hydrolase
MiR156 h	6.8E-5	Copper, zinc superoxide dismutase activity/zinc superoxide oxidoreductase
	3.9E-14	Amino acid derivative metabolism
	6.5E-10	Oxidoreductase activity/redox activity
	4.2E-8	Indole derivative metabolism
	7.8E-7	Response to wounding
MiR159b	1.2E-6	Response to stress
	1.3E-6	Response to external stimulus
	5.4E-6	Alternative oxidase activity
	7.9E-6	Oxidoreductase activity, acting on diphenols and related substances as donors, oxygen as acceptor
	2.5E-5	Oxidoreductase activity, acting on diphenols and related substances as donors
MiR160c	3.4E-5	Alternative oxidase activity
	5.0E-5	Oxidoreductase activity, acting on diphenols and related substances as donors, oxygen as acceptor
MiR165a	1.6E-4	Oxidoreductase activity, acting on diphenols and related substances as donors
	2.0E-6	Transcription factor activity
	7.0E-6	Transcription regulator activity
miR166f	4.8E-5	Flavonoid 3'-monooxygenase activity/flavonoid 3'-hydroxylase
	7.3E-5	DNA binding
	9.4E-5	Response to stimulus
	2.0E-8	Response to pathogen
	3.2E-8	Response to pest, pathogene or parasite
miR169d	3.8E-8	Response to external biotic stimulus
	9.5E-7	Response to biotic stimulus
	1.2E-6	Anthranilate synthase activity
	1.4E-6	Response to stress
	3.3E-6	Response to external stimulus
miR169j	4.2E-6	Oxo-acid-lyase activity
	2.7E-5	Anthranilate phosphoribosyltransferase activity
	3.4E-5	Ethylene biosynthesis/ethene biosynthesis
miR170a	3.4E-5	Ethylene metabolism/ethene metabolism
	8.2E-5	Amino acid derivative biosynthesis
	2.9E-5	Flavonoid 3'-monooxygenase activity/flavonoid 3'-hydroxylase
miR171a	5.0E-5	Transcription factor activity
	1.0E-4	Transcription regulator activity
	6.4E-5	Hydrolase activity, hydrolyzing O-glycosyl compounds/O-glucosyl hydrolase
miR172c	1.5E-4	Hydrolase activity, acting on glycosyl bonds/N-glycosylase/glycosylase
	1.2E-5	Alternative oxidase activity
	1.7E-5	Oxidoreductase activity, acting on diphenols and related substances as donors, oxygen as acceptor
miR172e	5.5E-5	Oxidoreductase activity, acting on diphenols and related substances as donors
	1.1E-5	Transcription regulator activity
	1.3E-5	Transcription factor activity
miR168a	1.6E-4	DNA binding
	1.9E-5	Beta-carotene hydroxylase
	6.8E-5	Activity carotene metabolism
miR395c	6.5E-20	Response to external stimulus
	1.1E-18	Response to stress
	1.1E-18	Response to stimulus
	6.9E-13	Response to abiotic stimulus
	1.1E-12	Response to biotic stimulus
miR395e	1.3E-12	Response to wounding
	4.8E-12	Catalytic activity/enzyme activity
	4.9E-11	Defense response/defence response
	4.5E-5	Copper, zinc superoxide dismutase activity/zinc superoxide oxidoreductase
	3.6E-6	Beta-carotene hydroxylase activity
miR395e	1.3E-5	Carotene metabolism
	1.5E-4	Tetraterpenoid metabolism
	1.5E-4	Carotenoid metabolism
	1.8E-4	Fucosyltransferase activity

Table IV Known light-related motifs in the upstream regions of miRNA genes that are predicted to be upregulated by UV-B

	1	2	3	4	5	6	7	8	9
mir156b	X	X	X	X			X		
mir156e	X	X	X					X	
mir156f	X	X	X				X		
mir156h	X	X							
mir157c	X	X	X			X			
mir159a	X	X		X					
mir159b	X	X	X	X	X	X			
mir160c	X	X	X						
mir165a	X	X		X	X	X			X
mir166c	X		X						
mir166f	X	X	X	X	X				
mir167d	X	X				X			
mir169d	X	X	X	X		X	X		
mir169j	X	X	X	X	X				
mir170	X	X	X				X		
mir171a	X	X	X	X	X				
mir172c	X	X	X	X		X			
mir172e	X	X	X				X		
mir393a	X	X	X	X	X	X			
mir398a	X	X	X						
mir401	X	X	X	X	X				
Total	21	20	17	11	7	7	5	1	1

1: GT-1 site, 2: I-box core (GATAA), 3: CCAAT-box (CCAAT), 4: TGA-box (TGACG), 5: TGACGT, 6: GATA-box (GATATTT), 7: I-box (GATAAGA), 8: G-box (CACGTG), 9: Hbox (CCTACC).

Table V Known stress-related motifs shared by UVB upregulated miRNA gene, miR167d, and its coregulated genes

miR167d	GATA-box	GT-1	I-box
AT1G20823	GATA-box	GT-1	TGA-box
AT4G16630	GATA-box		TGA-box
AT2G27830	GATA-box		
AT3G03270	GATA-box		
AT3G12510		GT-1	TGA-box

GATA-BOX: GATATTT, GT-1 site: GGTAA, I-box: GATAA, TGA-box: TGACG, CCAAT-box: CCAAT.

Expression-repression pathways that UV-B responsive miRNAs may be involved in

Light can trigger the transcription of a set of miRNA genes, which direct their target protein-coding mRNAs to be degraded quickly. Remarkably, eight of the 11 putative light-inducible miRNAs (except *miR393*, *miR398* and *miR401*) have targets that encode transcription factors. These targeted transcription factors can subsequently affect the expressions of their downstream genes. Hence, as in developmental stages, under light stress conditions there are downregulation pathways that are initiated by miRNAs, which cascade to the targets of these miRNAs, the targets of the targets of miRNAs, and so on. Table VI shows the targets of the putative UV-B responsive miRNAs.

A striking observation is that auxin signaling pathways can be affected by several light-induced miRNA genes. Auxin (principally indole-3-acetic acid) is an important hormone in plants. It affects many aspects of plant growth and development by influencing auxin response factors (ARF), a plant-

specific family of DNA binding proteins. ARFs regulate the expression of auxin-inducible genes such as GH3 and auxin/indole-3-acetic acid (Aux/IAA) by binding auxin response elements (AREs). Figure 2 illustrates possible auxin signaling pathways that four miRNAs, *miR160*, *miR165/166*, *miR167* and *miR393* may be involved in. As shown, these miRNAs will affect auxin signaling pathways by regulating different transcription factors under UV-B stimuli.

Materials and methods

Data sets

We studied all miRNA genes curated in the miRNA Registry (microrna.sanger.ac.uk/sequences/) as of January 1, 2006, except *miR408*, *miR413*, *miR414*, *miR419* and *miR420* that were not reported to be bonafide miRNAs (Xie *et al*, 2005). Three pairs of polycistronic miRNA genes, *miR169i* and *miR169j*, *miR169k* and *miR169l*, and *miR169m* and *miR169n* are referred to as *miR169j*, *miR169l* and *miR169n*, respectively, in this paper. Different from many animal miRNA genes, all these *A. thaliana* miRNAs have been annotated as intergenic genes, except *miR402* (Sunkar and Zhu, 2004). We studied a total of 109 miRNA genes.

The upstream sequences of protein-coding genes were downloaded from TAIR (ftp.arabidopsis.org/home/tair/seq_analysis_updates/). Known plant motifs were obtained from the motif database PLACE (www.dna.affrc.go.jp/PLACE/).

Microarray data for UV-B radiation from the international joint effort of *Arabidopsis* Gene Expression Project (AtGenExpress, www.arabidopsis.org/info/expression/ATGenExpress.jsp) were used in our study. The data were directly retrieved from TAIR site www.arabidopsis.org/info/expression/ATGenExpress.jsp. In these expression profiling experiments, 18-day-old *Arabidopsis* seedlings of Columbia-0 ecotype were harvested at the following time points: 0.25, 0.5, 1, 3, 6, 12 and 24 h after UV-B treatment. RNAs from root and shoot were analyzed separately using Affymetrix (www.affymetrix.com) Ath1 gene chip that contains more than 22K genes. Control samples were collected at the respective time points from plants grown under normal condition. To identify differentially expressed genes, we computed, for each gene at each time point profiled, the ratio of the expression level under UV-B treatment and the expression level under the control condition. A gene is considered upregulated if the gene expression ratio at any time point is at least 5, that is, we selected genes upregulated at least five-fold. This gave a total of 1280 genes upregulated in the root or shoot.

Outline of the computational approach

Our approach is designed to analyze the expression activities of plant miRNAs under a specified condition, by integrating a variety of DNA sequence data and gene expression data. As shown in Figure 3, the approach consists of four major steps:

Statistical analysis of target gene expression coherence: Under a particular condition, if an miRNA is upregulated, all its targets are very likely to be coherently downregulated if they are expressed. Thus, these target genes will have more similar expression profiles than a set of arbitrarily chosen genes. In this study, we used the average pairwise cosine similarity of gene expression (see section similarity of gene expression patterns) to measure the coherence of expression profiles of target genes for all reported miRNAs, and identified putative UV-B responsive miRNA families. For an miRNA family that has multiple member genes, we used the following steps to identify individual miRNA genes upregulated by UV-B stress.

Discovering cis-elements that are functionally relevant to UV-B stress: First we identified a comprehensive set of motifs from the 1280 protein-coding genes significantly upregulated under UV-B stress condition, using our WordSpy algorithm (Wang *et al*, 2005; Wang and Zhang, 2006) (see section motif identification). Then we applied a motif selection step to discover motifs that are functionally relevant to UV-B stress condition. If a motif is involved in the gene regulation

Table VI Target genes of UV-B responsive miRNAs

MiRNA	Target genes	Target gene functions	Binding motifs
miR156/157	At1g27360 (SPL1), At1g27370 (SPL10) At1g53160 (SPL4), At1g69170 (SPL4) At2g33810 (SPL3), At5g43270 (SPL2) At3g15270 (SPL5) At2g35320, At2g45990, At3g28690 At4g12080, At4g28660, At4g36860 At5g18590, At5g08620, At5g38610	Squamosa promoter Binding protein-like Transcription factor	TNCGTACAA
miR159/319	At3g11440, At4g18390, At5g06100 At3g06450, At3g61740, At3g15030 At4g37770, At5g09410, At4g26930 At5g17580, At5g56790, At5g55020 At2g26950, At5g67090, At2g32460 At1g30210, At1g53230, At2g26960 At2g31070	MYB factors	
miR160	At2g28350(ARF10), At4g30080(ARF16) At1g77850(ARF17)	Auxin response factors	
miR165/166	At1g52150 (ATHB-15), At2g34710 (PHB) At4g32880 (ATHB-8), At5g60690 (REV) At1g30490	HD-ZIP transcription factor	CAATNATTG
miR167	At5g37020, At5g19950 At1g02800, At3g61310	ARF transcription factor	
miR169	At1g17590, At1g54160, At1g72830 At3g20910, At5g12840 At1g70700, At1g80770	CBF HAP2-like factors	CCAAT
miR170/171	At4g00150, At5g61480 At2g45160, At3g60630	Scarecrow-like transcription factor	
miR172	At2g28550 (TOE1), At4g36920 (AP2) At5g60120 (TOE2), At5g67180 (TOE3) At2g39250 At2g42380 At5g65790 (MYB68) At3g14770 At5g36870 At3g44690, At5g12900, At5g19560	APETALA2-like factor APETALA2-like factor APETALA2-like factor bZIP family (G-box) MYB Nodulin MtN3 family protein Glycosyl transferase	CCACGTGG
miR393	At1g12820, At3g26810 At3g62980 (TIR1), At4g03190 At3g23690, At3g26830	F-box protein F-box protein BHLH	
miR398	At3g15640, At3g06370, At1g08050	Cytochrome <i>c</i> oxidase; zinc finger	
miR401	At2g13270, At3g42350		

under UV-B, most genes containing this motif should be upregulated under UV-B stress condition. For each motif reported by WordSpy, we assessed whether it was statistically significantly relevant to UV-B stress, with a *P*-value obtained from an accumulative hypergeometric test (see section hypergeometric distribution of motifs or motif combinations). Motifs with *P*-values less than 0.1 were selected to be used for identifying miRNA genes.

Locating the 5' proximal promoter regions of miRNA genes: To locate proximal promoters of miRNA genes, we first identified the transcription start sites or core promoter regions of the miRNA genes. In our study, we used the core promoters of 52 *Arabidopsis* miRNA genes experimentally identified by Carrington's laboratory (Xie *et al*, 2005), and the core promoters of the rest of miRNA genes predicted by our newly developed *de novo* core promoter prediction method called CoVote (Zhou *et al*, 2007). Using information of core promoters, we retrieved 1000 bp 5' proximal promoter regions of corresponding miRNA genes. For each miRNA gene, we scanned its proximal promoter region with the motifs discovered from previous step and obtained an array of motifs present in its promoter.

Identifying individual miRNA genes expressed under UV-B stress condition: Coexpressed genes are often coregulated, and coregulated genes contain common *cis*-elements or motif modules in their upstream regulatory regions. Therefore, the promoter regions of the genes responding to particular environmental stimuli must have characteristic *cis*-elements that are responsible for the upregulation of their expressions. If an miRNA gene *i* is upregulated under UV-B stress condition, the array of motifs present in its proximal promoter region are very likely to be involved in the upregulation of its expression. Most protein-coding genes that contain the same array of motifs should also be upregulated under the UV-B stress condition. For each miRNA gene, we assessed the statistical significance that the array of motifs contained in its promoters is relevant to UV-B stress, using a cumulative hypergeometric test (see section hypergeometric distribution of motifs or motif combinations). We selected miRNA genes with *P*-values less than 0.1 as candidate UV-B responsive genes. In order to increase accuracy, we further exploited the anticorrelation between the expression of an miRNA gene and the expressions of its targets to filter out possible false candidates. As we did not rely on any direct

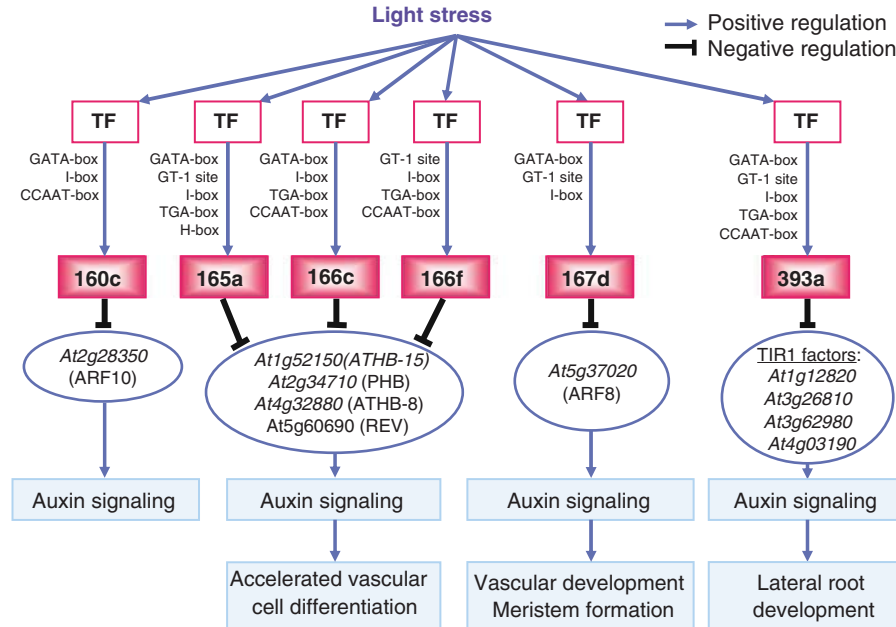


Figure 2 Auxin signaling pathways that putative UV-B responsive miRNAs may be involved in.

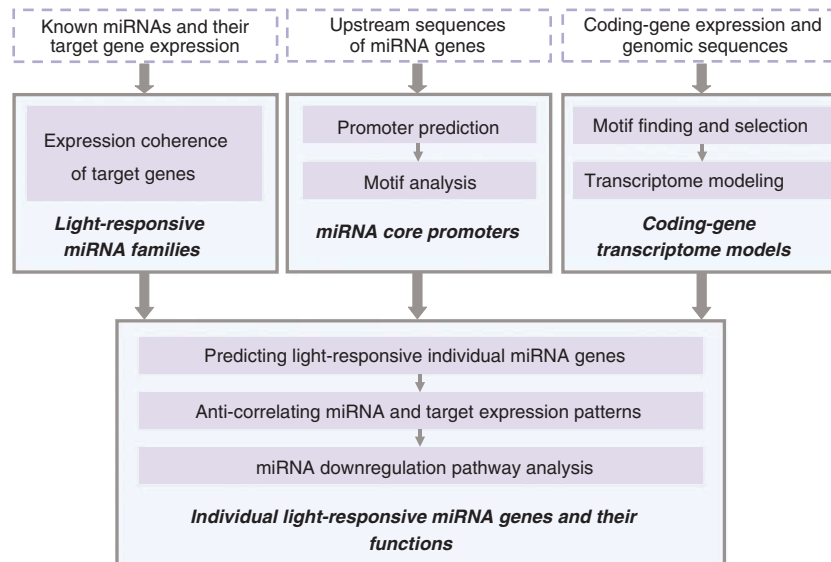


Figure 3 Sketch of the overall approach and four main components.

information of miRNA expression, we used the inferred expression of an miRNA gene and the expressions of its targets to compute their anticorrelation (see section inference of expression patterns of miRNA genes).

Motif identification

In the AtGenExpress microarray data, 1280 genes are upregulated at least five-fold in either shoot or root. The 1000 bp upstream promoters of these genes were analyzed in our study. We extracted statistically significant motifs, ranging from 5- to 9-mers, from the promoters of these genes using the WordSpy motif-finding algorithm (Wang *et al*, 2005; Wang and Zhang, 2006). WordSpy integrated statistical modeling and word counting methods, so that it is able to build a

dictionary of a large number of statistically significant motifs. WordSpy adopted a strategy of steganalysis, which is a technique for discovering hidden patterns and information from a media such as strings, so that it does not have to rely on additional background sequences and is still able to find motifs of nearly exact lengths. The details of the algorithm are available in (Wang *et al*, 2005; Wang and Zhang, 2006) and <http://cic.cs.wustl.edu/wordspy>.

Inference of expression patterns of miRNA genes

The rationale behind the inference was discussed in section outline of the computational approach. To reiterate, the method is transcriptome based; the main idea is to estimate the expected expression pattern of an miRNA using the expressions of putative coregulated protein-

coding genes, which are predicted based on the similarity of their upstream promoter regions with that of the miRNA. We used the significant *cis*-elements to measure the similarity of the promoters of an miRNA and a protein-coding gene. For two promoters, we first collected a set M of *cis*-elements that appeared in one or both of the promoters. We then represented a promoter as a vector, where each entry corresponds to a *cis*-element in M , with its value being 1 if it appears in the promoter, or 0 otherwise. The similarity of two promoters was then measured by the cosine similarity of their corresponding vectors (Menczer, 2004). Specifically, the similarity of the promoters of genes g and g_2 is defined as,

$$\text{similarity}(g_1, g_2) = \cos(v_1, v_2) = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|} \quad (1)$$

where $\vec{v}_1 \cdot \vec{v}_2$ is the inner product of motif vectors \vec{v}_1 and \vec{v}_2 of the two genes.

Using this cosine similarity, for each miRNA gene, we selected five protein-coding genes, whose promoters most closely resemble the promoter of the miRNA. We then used their average expression pattern as the inferred expression pattern of the miRNA. In our study, we also tested the Pearson correlation coefficient as similarity measure, and obtained similar result.

Hypergeometric distribution of motifs or motif combinations

The statistical significance of a motif, motif combination or motif module c was measured by a cumulative hypergeometric test (Altman, 1991). Given M genes on a microarray chip, assume that n of these M genes have a particular property (e.g. UV-B upregulated), m of the M genes on the chip contain the motif or motif combination c , and n of the N genes (which are UV-B upregulated) have the motif or motif module c . We calculated a P -value for the statistical significance of the motif or motif module c as the probability under which we would expect at least n genes to have c if we randomly selected m genes from the given M genes on the microarray chip. Specifically, P -value is computed as follows:

$$p(m, n, M, N) = \sum_{n \leq x \leq \min\{m, N\}} \frac{C_x^N C_{m-x}^{M-N}}{C_m^M}, \quad (2)$$

$$\text{where } C_m^M = \frac{M!}{m!(M-m)!}$$

Similarity of gene expression patterns

The similarity of expression patterns of two genes was measured by cosine similarity defined as in Equation (1), where $\vec{v}_1 \cdot \vec{v}_2$ is the inner product of expression profile vectors \vec{v}_1 and \vec{v}_2 of the two genes. In our study, we also used the Pearson correlation coefficient to measure the similarity of expression patterns of two genes, and obtained similar result (data not shown).

Acknowledgements

We thank Jianhua Ruan for stimulating discussions. This work was supported in part by NSF grants ITR/EIA-0113618 and IIS-0535257.

References

- Altman D (1991) *Practical Statistics for Medical Research*. Chapman & Hall/CRC, New York, USA
- Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol* **55**: 373–399
- Arguello-Astorga G, Herrera-Estrella L (1998) Evolution of light-regulated plant promoters. *Annu Rev Plant Physiol Plant Mol Biol* **49**: 525–555

- Aukerman M, Sakai H (2003) Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell* **15**: 2730–2741
- Axtell M, Bartel D (2005) Antiquity of microRNAs and their targets in land plants. *Plant Cell* **17**: 1658–1673
- Bari R, Pant BD, Stitt M, Scheible W (2006) PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiol* **141**: 988–999
- Baroli I, Gutman B, Ledford H, Shin J, Chin B, Havaux M, Niyogi K (2004) Photo-oxidative stress in a xanthophyll-deficient mutant of *Chlamydomonas*. *J Biol Chem* **279**: 6337–6344
- Bartel D (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**: 281–297
- Beer M, Tavazoie S (2004) Predicting gene expression from sequence. *Cell* **117**: 185–198
- Bonnet E, Wuyts J, Rouze P, de Peer YV (2004) Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proc Natl Acad Sci USA* **101**: 11511–11516
- Bussemaker H, Li H, Siggia E (2001) Regulatory element detection using correlation with expression. *Nat Genet* **27**: 167–171
- Chattopadhyay S, Puente P, Deng X, Wei N (1998) Combinatorial interaction of lightresponsive elements plays a critical role in determining the response characteristics of lightregulated promoters in *Arabidopsis*. *Plant J* **15**: 69–77
- Chen X (2004) A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science* **303**: 2022–2025
- Chiou TJ, Aung K, Lin SI, Wu CC, Chiang SF, I Su C (2006) Regulation of phosphate homeostasis by microRNA in *Arabidopsis*. *Plant Cell* **18**: 412–421
- Chory J, Chatterjee M, Cook R, Elich T, Fankhauser C, Li J, Nagpal P, Neff M, Pepper A, Poole D, Reed J, Vitart V (1996) From seed germination to flowering, light controls plant development via the pigment phytochrome. *Proc Natl Acad Sci USA* **93**: 12066–12071
- Dunaeva M, Adamska I (2001) Identification of genes expressed in response to light stress in leaves of *Arabidopsis thaliana* using RNA differential display. *Eur J Biochem* **268**: 5521–5529
- Eshed Y, Baum S, Perea J, Bowman J (2001) Establishment of polarity in lateral organs of plants. *Curr Biol* **11**: 1251–1260
- Fujii H, Chiou T, Zhu J (2005) A microRNA involved in phosphate starvation response in *Arabidopsis*. *Curr Biol* **15**: 2038–2043
- Harvaux M, Kloppstech K (2001) The protective functions of carotenoid and flavonoid pigments against excess visible radiation at chilling temperature investigated in *Arabidopsis* npq and tt mutants. *Planta* **213**: 953–966
- He L, Hannon G (2004) MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* **5**: 522–531
- Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant *cis*-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Res* **27**: 297–300
- Houbaviy H, Dennis L, Jaenisch R, Sharp P (2005) Characterization of a highly variable eutherian microRNA gene. *RNA* **11**: 1245–1257
- Jiao Y, Ma L, Strickland E, Deng X (2005) Conservation and divergence of light-regulated genome expression patterns during seedling development in rice and *Arabidopsis*. *Plant Cell* **17**: 3239–3256
- Jones-Rhoades M, Bartel D (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell* **14**: 787–799
- Kidner C, Martienssen R (2004) Spatially restricted microRNA directs leaf polarity through ARGONAUTE1. *Nature* **428**: 81–84
- Kimura M, Manabe K, Abe T, Yoshida S, Matsui M, Yamamoto Y (2003a) Analysis of hydrogen peroxide-independent expression of the. *Photochem Photobiol* **77**: 668–674
- Kimura M, Yamamoto Y, Seki M, Sakurai T, Sato M, Abe T, Yoshida S, Manabe K, Shinozaki K, Matsui M (2003b) Identification of *Arabidopsis* genes regulated by high light-stress using cDNA microarray. *Photochem Photobiol* **77**: 226–233
- Kimura M, Yoshizumi T, Manabe K, Yamamoto Y, Matsui M (2001) *Arabidopsis* transcriptional regulation by light stress via hydrogen

- peroxide-dependent and -independent pathways. *Genes Cells* **6**: 607–617
- Lee J, Colinas J, Wang J, Mace D, Ohler U, Benfey P (2006) Transcriptional and posttranscriptional regulation of transcription factor expression in *Arabidopsis* roots. *Proc Natl Acad Sci USA* **103**: 6055–6060
- Lee Y, Kim M, Han J, Yeom K, Lee S, Baek S, Kim V (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**: 4051–4060
- Llave C, Xie Z, Kasschau K, Carrington J (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* **297**: 2053–2056
- Lu S, Sun Y, Shi R, Clark C, Li L, Chiang V (2005) Novel and mechanical stress-responsive MicroRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. *Plant Cell* **17**: 2186–2203
- McConnell J, Emery J, Eshed Y, Bao N, Bowman J, Barton M (2001) Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. *Nature* **411**: 709–713
- Menczer F (2004) Evolution of document networks. *Proc Natl Acad Sci USA* **101** (Suppl 1): 5261–5265
- Narusaka Y, Narusaka M, Seki M, Umezawa T, Ishida J, Nakajima M, Enju A, Shinozaki K (2004) Crosstalk in the responses to abiotic and biotic stresses in *Arabidopsis*: analysis of gene expression in cytochrome P450 gene superfamily by cDNA microarray. *Plant Mol Biol* **55**: 327–342
- Palatnik J, Allen E, Wu X, Schommer C, Schwab R, Carrington J, Weigel D (2003) Control of leaf morphogenesis by microRNAs. *Nature* **425**: 257–263
- Schwab R, Palatnik J, Riester M, Schommer C, Schmid M, Weigel D (2005) Specific effects of microRNAs on the plant transcriptome. *Dev Cell* **8**: 517–527
- Shao N, Vallon O, Dent R, Niyogi K, Beck C (2006) Defects in the cytochrome *b6/f* complex prevent light-induced expression of nuclear genes involved in chlorophyll biosynthesis. *Plant Physiol* **141**: 1128–1137
- Shinozaki K, Yamaguchi-Shinozaki K, Seki M (2003) Regulatory network of gene expression in the drought and cold stress responses. *Curr Opin Plant Biol* **6**: 410–417
- Sunkar R, Zhu J (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell* **16**: 2001–2019
- Teakle G, Kay S (1995) The GATA-binding protein CGF-1 is closely related to GT-1. *Plant Mol Biol* **29**: 1253–1266
- Terzaghi W, Cashmore A (1995) Light-regulated transcription. *Ann Rev Plant Physiol Plant Mol Biol* **46**: 445–474
- Vaucheret H, Vazquez F, Crete P, Bartel D (2004) The action of ARGONAUTE 1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev* **18**: 1187–1197
- Venter M, Botha F (2004) Promoter analysis and transcription profiling: integration of genetic data enhances understanding of gene expression. *Physiol Plant* **120**: 74–83
- Wang G, Yu T, Zhang W (2005) WordSpy: identifying transcription factor binding motifs by building a dictionary and learning a grammar. *Nucleic Acids Res* **33**: W412–W416
- Wang G, Zhang W (2006) A steganalysis-based approach to comprehensive identification and characterization of functional regulatory elements. *Genome Biol* **7**: R49
- Wellmer F, Riechmann J (2005) Gene network analysis in plant development by genomic technologies. *Int J Dev Biol* **49**: 745–759
- Xie Z, Allen E, Fahlgren N, Calamar A, Givan S, Carrington J (2005) Expression of *Arabidopsis* MIRNA genes. *Plant Physiol* **138**: 2145–2154
- Xie Z, Kasschau K, Carrington J (2003) Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation. *Curr Biol* **13**: 784–789
- Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of *cis*-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends Plant Sci* **10**: 88–94
- Zhang J, Creelman R, Zhu J (2004) From laboratory to field. Using information from *Arabidopsis* to engineer salt, cold, and drought tolerance in crops. *Plant Physiol* **135**: 615–621
- Zhou X, Ruan J, Wang G, Zhang W (2007) Characterization and identification of MicroRNA promoters in four model species. *PLoS Computat Biol* (in press) Provisional doi:10.1371/journal.pcbi.0030037.eor